From the Division of Medical Imaging and Technology
Department of Clinical Science, Intervention and Technology
Karolinska Institutet, Stockholm, Sweden

MULTIVARIATE DATA ANALYSIS
APPLIED TO MRS AND MRI STUDIES
OF AGING AND SPINAL CORD INJURY

Johanna Öberg

Stockholm 2011
To my family
Abstract

Magnetic resonance can be used for non-invasive studies of the body without the use of ionizing radiation. Magnetic resonance imaging and magnetic resonance spectroscopy have proven to be valuable utilities for research in life sciences. This thesis deals with magnetic resonance investigations of the central nervous system in vivo and is based on four studies. In studies I-III in vivo proton magnetic resonance spectroscopy data were acquired in three animal models. These models were designed to monitor Alzheimer’s disease, spinal cord injury and premature aging. We wanted to quantify and evaluate the differences in metabolite levels in diseased animals in comparison with controls. In study IV, resting-state functional magnetic resonance imaging was applied to investigate young and elderly human subjects. Three different pre-processing procedures were also evaluated. Furthermore, in this thesis we aimed to explore how data acquired with magnetic resonance spectroscopy and functional magnetic resonance imaging can be extracted and analyzed using model free and model driven multivariate data analyses. The linear multivariate data analysis methods principal components analysis and partial least squares projections to latent structures were applied to magnetic resonance spectroscopy data acquired in rodents. Independent component analysis was applied to the resting-state functional magnetic resonance imaging data acquired in human subjects.

Group differences in brain metabolites between diseased and control animals were observed and reported in study I-III. By applying the method partial least squares projections to latent structures to all detected metabolites, we were able to develop models that could differentiate the diseased rodents from the normal controls and evaluate the sensitivity and specificity of the models.

In study IV we investigated the effects of preprocessing prior to independent component analysis. We found that global signal removal can enhance anti-correlation in resting-state functional connectivity networks. We also found that normal brain aging can lead to significant changes in functional connectivity.
List of Publications

I  Age related changes in brain metabolites observed by $^1$H MRS in APP/PS1 mice

II  $^1$H MRS in spinal cord injury: acute and chronic alterations in rat brain and lumbar spinal cord
   Erschbamer M,* Oberg J,* Westman E, Sitnikov R, Olson L, Spenger C

III High brain lactate is an early hallmark of aging and due to a shift in LDH-A/LDH-B ratio

IV Aging related changes in functional connectivity networks investigated by resting-state BOLD functional MRI: a study of independent component analysis and image preprocessing procedures
   Oberg J, Jonsson T, Li T-Q
   Submitted manuscript.

*Equal contribution
# Contents

1 Introduction ................................................................................. 1
   1.1 Magnetic resonance ......................................................... 1
      1.1.1 Magnetic resonance spectroscopy ............................... 2
      1.1.2 Functional magnetic resonance imaging ................. 7
   1.2 Multivariate data analysis .................................................. 8
      1.2.1 Principal components analysis .................................. 8
      1.2.2 Partial least squares projections to latent structures ...... 12
      1.2.3 Independent components analysis ............................. 14
   1.3 Overview of investigated diseases and conditions .............. 16
      1.3.1 Alzheimer’s disease ............................................... 16
      1.3.2 Spinal cord injury .................................................. 17
      1.3.3 Aging ................................................................. 17
2 Aim of thesis ............................................................................. 19
3 Materials and methods ............................................................. 21
   3.1 Magnetic resonance spectroscopy ...................................... 21
      3.1.1 Study I ................................................................. 22
      3.1.2 Study II ............................................................... 23
      3.1.3 Study III ............................................................ 24
   3.2 Functional magnetic resonance imaging ............................. 24
      3.2.1 Study IV .............................................................. 24
4 Results ..................................................................................... 27
   4.1 Study I .............................................................................. 27
   4.2 Study II ............................................................................. 29
   4.3 Study III ............................................................................ 31
   4.4 Study IV ............................................................................ 32
5 Discussion ................................................................................ 35
6 Acknowledgements ................................................................... 43
Bibliography ................................................................................ 45
1. Introduction

1.1 Magnetic resonance

Magnetic resonance (MR) can be used for non-invasive studies of the body without the use of ionizing radiation. Magnetic resonance imaging (MRI) is the most widely known application of MR. MRI provides information about the three-dimensional structure of objects and is therefore useful to examine brain anatomy and pathology. In addition to obtaining structural information, functional MRI (fMRI) can be used to investigate brain function. Magnetic resonance spectroscopy (MRS), on the other hand, provides information about the chemical constituents of objects, including information on tissue metabolism and neurotransmitters. Common to the techniques mentioned above is that the object under investigation is placed in a static magnetic field, $B_0$.

When placing a macroscopic sample in the magnetic field, the $^1$H nuclei (protons) in the sample will be affected and cause a net magnetization aligned with $B_0$. By applying an excitation pulse, in the radio frequency (RF) range, absorption and emission of electromagnetic radiation from the sample can be observed. The pulse needs to have a bandwidth around a specific resonance frequency, the Larmor frequency ($v_0$), to satisfy the Larmor equation

$$v_0 = \gamma B_0. \quad (1.1)$$

After the excitation pulse has been applied, a small fraction of the proton spins in the sample will be excited and eventually return to ground state by emitting the absorbed energy, which is the detected MR signal. To create MR images the signal needs to be spatially encoded, which is achieved using magnetic field gradients; weak magnetic fields that change linearly with position and are superimposed on $B_0$. Gradients are also used to localize the volume of interest (VOI) when performing MRS. To carry out the MR acquisition of interest, RF pulses, gradient pulses and the timing of the data acquisition are controlled by a program, the pulse sequence.
1.1.1 Magnetic resonance spectroscopy

Chemical shift

All protons in a sample do not share the same resonance frequency. This is due to that the resonance frequency of a proton in a magnetic field is not only dependent on \( \gamma \) and \( B_0 \) as in Eq. 1.1. Protons in different molecules, and within the same molecule, absorb energy of different frequencies, which is referred to as the chemical shift (\( \delta \)). Chemical shift is caused by shielding from the electrons surrounding the proton. The effect is that the proton may experience a magnetic field, \( B \), different from \( B_0 \)

\[
B = B_0 (1 - \sigma), \quad (1.2)
\]

where \( \sigma \) is the magnetic shielding constant. As a consequence the proton may also have a resonance frequency different from \( \nu_0 \). The chemical shift is expressed in parts per million (ppm) of the resonance frequency of a reference molecule and is by convention defined as

\[
\delta = \frac{\nu - \nu_{\text{ref}}}{\nu_{\text{ref}}} \times 10^6, \quad (1.3)
\]

where \( \nu \) is the resonance frequency of the magnetically equivalent protons, and \( \nu_{\text{ref}} \) is a reference frequency. Expressed in this form the \( B_0 \) dependence is removed. Hence, referring to the chemical shift of a molecule, or a part of a molecule, is common to MRS experiments carried out with magnets of different \( B_0 \) fields.

The MRS signal

Figure 1.1 shows an example of a volume from which it might be interesting to acquire an MR spectrum: a cubic VOI of eight mm\(^3\) in the dorsal hippocampus of a mouse brain. The purpose of a volume selective spectroscopy pulse sequence is to localize, and acquire signal from, the VOI.

Using the pulse sequence of point resolved spectroscopy (PRESS) [1], one slice selective excitation pulse is followed by two slice selective refocusing pulses. Gradients along the three orthogonal axes are applied in the presence of each of the RF pulses to achieve three-dimensional localization of MR signals persisting to the double refocused echo.

The signal is acquired using a receiver coil and is typically a superposition of signals of different frequencies, amplitudes, phases and decays (Figure 1.2 (left)). The signal is sampled, digitized and stored in a computer for further processing. Conversion of time domain-data to frequency-domain data by Fourier transformation will reveal the resonances that are present in the signal as peaks in a spectrum. Figure 1.2 (right) shows the signal to the left in the same figure after Fourier transformation, with the relative frequencies in ppm on the horizontal axis. Molecular groups generate specific resonance patterns in the spectrum, either as single peaks, doublets or more complex
Figure 1.1: Illustration of a VOI in the dorsal hippocampus of a mouse brain. The VOI, an eight mm$^3$ cube, is shown with a cut-out on a scull-stripped brain template.

spectra. The signal originating from water (centered at 4.7 ppm), if not suppressed, is often more than 10 000 times the signal of the molecules of interest. Water-suppression can provide a more reliable and consistent detection of these molecules.

Figure 1.2: Detected MRS signal presented in the time domain (left) and in the frequency domain after Fourier transformation (right). The largest peak in the right panel corresponds to water.

Figure 1.3 shows the small peaks to the right of the water peak in Figure 1.2 (right), ’zoomed in’. This part of the in vivo $^1$H MR spectrum was the focus of the MRS studies in this thesis. The spectrum was acquired from a voxel like the one shown in Figure 1.1, and the peaks in the spectrum correspond to metabolites present in hippocampus of a mouse brain. Metabolites are products of metabolism that originate from chemical reactions in the body. The
peaks in Figure 1.3 are marked with the metabolite abbreviations presented in Table 1.1.

A high magnetic field homogeneity will enable close peaks in the spectrum to be distinguished. Prior to applying the PRESS sequence the magnetic field homogeneity is therefore optimized locally in the VOI. This procedure is referred to as localized shimming.

*Figure 1.3:* An *in vivo* proton MR spectrum acquired from hippocampus in a mouse brain with a 9.4 T scanner.
<table>
<thead>
<tr>
<th>Metabolites used in basis set in study I-III&lt;sup&gt;(a)&lt;/sup&gt;</th>
<th>Abbreviation used in thesis</th>
<th>Study in which metabolite was regarded as detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetate</td>
<td>✓</td>
<td>I, II&lt;sup&gt;(c)&lt;/sup&gt;, III</td>
</tr>
<tr>
<td>N-acetyl Aspartate</td>
<td>✓</td>
<td>I, II&lt;sup&gt;(c)&lt;/sup&gt;, III</td>
</tr>
<tr>
<td>N-acetyl Aspartyl Glutamate</td>
<td>✓</td>
<td>I, II&lt;sup&gt;(c)&lt;/sup&gt;, III</td>
</tr>
<tr>
<td>Adenosine Triphosphate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>✓</td>
<td>I, II, III</td>
</tr>
<tr>
<td>γ-Aminobutyric Acid</td>
<td>✓</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Ascorbic Acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartate</td>
<td>✓</td>
<td>I</td>
</tr>
<tr>
<td>Choline-containing Compounds</td>
<td>✓&lt;sup&gt;(b)&lt;/sup&gt;</td>
<td>I, II</td>
</tr>
<tr>
<td>Creatine and Phosphocreatine</td>
<td>✓</td>
<td>I, II</td>
</tr>
<tr>
<td>Glucose</td>
<td>✓</td>
<td>I, II</td>
</tr>
<tr>
<td>Glutamate</td>
<td>✓</td>
<td>I, II</td>
</tr>
<tr>
<td>Glutamine</td>
<td>✓</td>
<td>I, II</td>
</tr>
<tr>
<td>Glutathione</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyceral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycogen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histidine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homocarnosine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-Hydroxybutyrate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myo-Insitol</td>
<td>✓</td>
<td>I, II</td>
</tr>
<tr>
<td>Scyllo-Insitol</td>
<td>✓</td>
<td>III</td>
</tr>
<tr>
<td>Lactate</td>
<td>✓</td>
<td>III</td>
</tr>
<tr>
<td>Macromolecules</td>
<td>✓</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyruvate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succinate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taurine</td>
<td>✓</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Threonine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tryptophan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intra- and Extramyocellular lipids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deoxymyoglobin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citric Acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carnosine</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Guanidinoacetate (Gua) was also included in the basis set, and regarded as detected in study I and II.
(b) Glycerophosphocholine (GPC) and Phosphocholine (PCh).
(c) Aspartate was included in the basis set, and regarded as detected in study I and II.
(d) The sum Cr + PCr is abbreviated as tCr.
Spectral quantification

The total MRS signal obtained from a mixture of compounds can be seen as a linear combination of the signals from the isolated compounds [2]. The program LCModel [3] fits a linear combination of model spectra (the basis set) to the \textit{in vivo} spectrum of interest. Linear combination modeling algorithms essentially adjust the amplitudes, frequencies, line widths and phases of the metabolite basis set to match the \textit{in vivo} spectrum as close as possible [2]. An example of the output of an LCModel analysis is shown in Figure 1.4. The spectrum is plotted as a thin curve and the thick red curve is the LCModel fit to the data. Also plotted as a thin curve, below the spectrum, is the baseline. At the top the residuals, i.e. the data minus the fit to the data, are shown. The residuals are a sensitive diagnostic of the analysis and should appear randomly scattered about zero.

For a given noise level the lowest possible quantification errors are given by the Cramér-Rao lower bounds (CRLBs). CRLBs increase with increasing spectral overlap. The table to the right in Figure 1.4 summarizes the quantified metabolite concentrations together with the corresponding CRLBs. CRLBs are an objective quality control that account for both resolution and noise level, and can be used as a guide to the reliability of the estimates. A commonly used approach is to only consider quantified metabolites with corresponding CRLBs < 20 % [4]. Others have used a CRLB limit of 50 % [5]. In this work, limits of CRLB < 50 % (study I and III) and CRLB < 20 % (study II) were used.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Example of output of LCModel quantification with frequency-domain data plotted as a thin curve. The red curve corresponds to the LCModel fit to the data. At the top the residuals are shown. A concentration table to the right summarizes the quantified metabolites together with corresponding CRLBs.}
\end{figure}
1.1.2 Functional magnetic resonance imaging

fMRI data are acquired as a time series of images in which contrast changes over time are investigated. fMRI based on the blood-oxygen-level dependent (BOLD) contrast relies on the fact that oxygenated and deoxygenated hemoglobin in the blood have different magnetic properties. Block design is a commonly used experimental setup in fMRI, in which the subject is instructed to perform experimental and control tasks in an alternating sequence of blocks. Biswal 1995 et al. [6] were the first to report resting-state (i.e. not related to a specific task) functional connectivity networks in the brain. They found temporal correlation across functionally related areas in spontaneous low-frequency fluctuations of resting-state BOLD fMRI signals. Functional connectivity magnetic resonance imaging (fcMRI), takes the advantage of the similarity in large-amplitude low-frequency fluctuations of the BOLD signal intensity across spatially separated brain regions [7].

Common to the different analysis methods used for resting-state fMRI is the preprocessing procedure of the fMRI data employed to enhance the quality of the functional connectivity mapping. However, the preprocessing steps used in the literature are far from standardized, which makes it difficult to directly compare the functional connectivity results from different studies of the same neural pathology. For example, many resting-state and task-based fMRI studies included some type of correction for the global signal averaged across the entire brain. This is thought to enhance the observation of localized neuronal effects. One reproducible consequence of the global signal removal has been the observation of increased anti-correlation in a number of functional brain networks.
1.2 Multivariate data analysis

Multivariate data analysis (MVDA) is a collective name for a number of statistical methods for exploration and analysis of multidimensional data. To find a suitable representation of a complex dataset, a principle of optimal transform is usually defined to represent the original data. Examples of principles are dimension reduction, statistical ‘interestingness’ of the resulting components, simplicity of the transformation, or other criteria, including application-oriented ones [13].

In this thesis, the linear MVDA methods principal components analysis (PCA), partial least squares projections to latent structures (PLS) and independent component analysis (ICA) were applied (Figure 1.5). PCA and PLS are both projection methods that convert multi-dimensional data into lower-dimensional representations. A principal component analysis is performed in both PCA and PLS but they rarely result in the exact same principal components, since their aims are different. PCA finds the direction of maximum variance in the data matrix. PLS, on the other hand, works to find the best correlation between the data matrix and another matrix. ICA is a transform in which the desired representation is the one that minimizes the statistical dependence of the representation. PCA and ICA are both unsupervised (model-free) methods, whereas PLS is a supervised (model-driven) method.

![Multivariate data analysis in thesis](image)

*Figure 1.5:* Overview of the multivariate data analysis methods that were used for analysis of MRS and fMRI data in this thesis: principal components analysis (PCA), partial least squares projections to latent structures (PLS) and independent component analysis (ICA).

1.2.1 Principal components analysis

PCA was first formulated in statistics by Karl Pearson in 1901 [14]. Example data, designed by Stefan Rännar [15], will be used here to illustrate the principle of PCA and PLS. Ten women and ten men are asked about three things concerning themselves: shoe size, weight and length. Their answers
are collected in a matrix, $X$ (Table 1.2\textsuperscript{1}). Looking at one variable at a time it is not possible to separate women from men since there are overlaps in the data, i.e. some women are taller than some men etc. The question is then: can PCA, based on all variables in $X$, separate women from men? The multivariate analysis of the example data was carried out using [16].

Table 1.2: The answers from the survey of ten women and ten men regarding shoe size, length and weight.

<table>
<thead>
<tr>
<th>Sex</th>
<th>Shoe size</th>
<th>Length [cm]</th>
<th>Weight [kg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>37</td>
<td>167</td>
<td>57</td>
</tr>
<tr>
<td>Female</td>
<td>36</td>
<td>170</td>
<td>54</td>
</tr>
<tr>
<td>Male</td>
<td>42</td>
<td>167</td>
<td>71</td>
</tr>
<tr>
<td>Female</td>
<td>40</td>
<td>173</td>
<td>62</td>
</tr>
<tr>
<td>Male</td>
<td>42</td>
<td>174</td>
<td>76</td>
</tr>
<tr>
<td>Male</td>
<td>44</td>
<td>181</td>
<td>78</td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>175</td>
<td>53</td>
</tr>
<tr>
<td>Female</td>
<td>35</td>
<td>165</td>
<td>51</td>
</tr>
<tr>
<td>Male</td>
<td>40</td>
<td>189</td>
<td>71</td>
</tr>
<tr>
<td>Male</td>
<td>44</td>
<td>178</td>
<td>73</td>
</tr>
<tr>
<td>Female</td>
<td>41</td>
<td>168</td>
<td>51</td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>174</td>
<td>49</td>
</tr>
<tr>
<td>Male</td>
<td>42</td>
<td>174</td>
<td>62</td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>162</td>
<td>62</td>
</tr>
<tr>
<td>Male</td>
<td>40</td>
<td>184</td>
<td>78</td>
</tr>
<tr>
<td>Male</td>
<td>41</td>
<td>181</td>
<td>81</td>
</tr>
<tr>
<td>Female</td>
<td>37</td>
<td>168</td>
<td>50</td>
</tr>
<tr>
<td>Female</td>
<td>39</td>
<td>172</td>
<td>53</td>
</tr>
<tr>
<td>Male</td>
<td>44</td>
<td>175</td>
<td>78</td>
</tr>
<tr>
<td>Male</td>
<td>42</td>
<td>182</td>
<td>78</td>
</tr>
</tbody>
</table>

The observations can be represented as points in a multidimensional space where the variables define the axes (Figure 1.6 (left)). The lengths of the axes are determined by the scaling of the variables. If one has no prior information about the importance of the variables, autoscaling all variables to unit variance is recommended [17]. To the right in Figure 1.6 the points are shown after scaling to unit variance and mean centering.

In Figure 1.6 (right) the first ($t[1]$) and second ($t[2]$) principal components that are the result of a PCA based on $X$ are also plotted. $t[1]$ best approximates the data in a least squares sense and represents the maximum variance

\textsuperscript{1}It should be noted that 7/10 women (and none of the men) are underweight in this example and their assumed body shapes do not reflect reality very well, at least not in the developed countries. However, the example data are still useful to explain PCA.
direction in the data. $t[2]$ is computed in a direction that is orthogonal to $t[1]$ and reflecting the second largest source of variation. Projection of the points onto the plane spanned by the two first principal components is presented in Figure 1.7. This is shown in the $t[1]/t[2]$ score plot (upper), which is a summary of the relationships among the observations (women and men). A corresponding loading plot (lower) is a summary of the variables (shoe size, length and weight). The two plots are superimposable: a direction in one plot correspond to the same direction in the other plot. A point in the score plot, a score, represent one individual and all three variables that were registered for that individual. The confidence ellipse is based on Hotelling’s $T^2$, a multivariate generalization of Student’s t-test, at significance level 0.05. $T^2$ measures how far away an observation is from the center of the model and provides a tolerance region for the data in a two-dimensional score plot. The loadings can be used to interpret the score. The position of an observation in a given direction in a score plot is influenced by variables lying in the same direction in the loading plot. Hence all three variables length, weight and shoe size were important for the separation of men and women in the score scatter plot in Figure 1.7.

In this example, a two-dimensional representation of the original three-dimensional data was formed. The method can be extended to create two-dimensional representations of data that originally consisted of many more variables. Looking at the score plot we see that the data points separate into the two groups men and women, so to answer the original question: yes, PCA could separate women from men using the data in Table 1.2.
To evaluate the performance of the model one can look at how well the model fitted the data, and how good the predictive power of the model is. The distance from the observations in variable space to the model plane in Figure 1.6 represent the amount of variation that is unexplained by the model, the residuals. Hence $X$ is approximated by a least squares plane, spanned by $t[1]$ and $t[2]$, and the residuals. The explained variation, $R^2$, is a measure of how well the model was able to fit the original data. $R^2$, ranging between 0 and 1, is defined as

$$R^2 = 1 - \frac{\text{residual sum of squares}}{\text{total variation in } X \text{ after mean centering}}.$$  \hfill (1.4)

Cross-validation has become standard in MVDA to test the significance of a PC- or a PLS-model [18]. Cross-validation means that a part of the data is left out, and used as a prediction set in the model. The procedure is repeated in a systematic way until all the data have been left out and predicted. The sum of squared differences between predicted and observed data serve as a measure
of the predictive power of the model. $Q^2$ (ranging between 0 and 1) specifies the predicted variation and is calculated as

$$Q^2 = 1 - \frac{\text{predictive residual sum of squares}}{\text{total variation in } X \text{ after mean centering}}. \quad (1.5)$$

If a new principal component enhances the predictive power compared with the preceding principal component, the new component is kept in the model. If not, the new principal component is said to be insignificant, and no more components are calculated for the model. In the PC-model presented in Figure 1.7 there was one significant component, $t[1]$. (The second component, $t[2]$, was computed only to enable a two-dimensional visualization of the data.) There are a number of rules in [16] that are used to decide if a component is significant or not. In the PC-model in Figure 1.7 the limit was set to $Q^2 > 0.29$. $R^2$ and $Q^2$ for $t[1]$ were 0.76 and 0.42 respectively. In conclusion, PCA of the data in Table 1.2 resulted in one significant principal component with an explained variation of 76% ($R^2 = 0.76$) and a predicted variation of 42% ($Q^2 = 0.42$).

As PCA was applied to MRS data quantified with LCModel in this thesis the variables correspond to the different metabolites. Hence, the number of variables in the analysis depends on the method of quantification. For analysis of MRS data PCA can be applied to get an overview of the data, and track outliers, before building PLS-models for classifications. PCA can also be used as a preprocessing tool of fMRI data prior to ICA as in study IV. PCA is then performed to reduce the dimension of the data so that the maximum amount of variance is preserved.

### 1.2.2 Partial least squares projections to latent structures

PLS, a regression extension of PCA, is a method for relating two data matrices to each other by a linear multivariate model [19, 20]. Many studies typically constitute of a set of controls and treated/diseased samples or subjects. Sometimes additional knowledge of the samples is also of interest, e.g. dose, age, gender and diet. The additional information can be put in another matrix, $Y$. PLS can then be applied to the data with the aim to predict $Y$ from $X$. PLS forms ‘new X-variables’, principal components, as linear combinations of the original variables, and thereafter uses the components as predictors of $Y$ [18]. $Y$ may contain both quantitative (e.g. dose, age, concentration) and qualitative (e.g. control/diseased) data. In the case of qualitative data in $Y$, PLS analysis is referred to as PLS-discriminant analysis (PLS-DA) [21, 22]. $Y$ then encodes class membership by a set of ‘dummy’ variables (e.g. zeros and ones). The dummy matrix, $Y$, for $X$ in Table 1.2, is presented Table 1.3. $Y$ has two dummy variables ($Y_1$ and $Y_2$), one for each class (women and men).
Table 1.3: Dummy matrix $Y$ ($Y_1$ and $Y_2$) for PLS-DA of the data in Table 1.2. Also shown are the predictions of class (Predicted $Y_1$ and Predicted $Y_2$). Using a cut-off of 0.5 all samples will be considered correctly classified.

<table>
<thead>
<tr>
<th>Identification</th>
<th>$Y_1$</th>
<th>$Y_2$</th>
<th>Predicted $Y_1$</th>
<th>Predicted $Y_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female 1</td>
<td>1</td>
<td>0</td>
<td>0.99</td>
<td>0.01</td>
</tr>
<tr>
<td>Female 2</td>
<td>1</td>
<td>0</td>
<td>1.04</td>
<td>-0.04</td>
</tr>
<tr>
<td>Female 3</td>
<td>1</td>
<td>0</td>
<td>0.56</td>
<td>0.44</td>
</tr>
<tr>
<td>Female 4</td>
<td>1</td>
<td>0</td>
<td>0.81</td>
<td>0.19</td>
</tr>
<tr>
<td>Female 5</td>
<td>1</td>
<td>0</td>
<td>1.27</td>
<td>-0.27</td>
</tr>
<tr>
<td>Female 6</td>
<td>1</td>
<td>0</td>
<td>0.80</td>
<td>0.20</td>
</tr>
<tr>
<td>Female 7</td>
<td>1</td>
<td>0</td>
<td>0.91</td>
<td>0.09</td>
</tr>
<tr>
<td>Female 8</td>
<td>1</td>
<td>0</td>
<td>0.94</td>
<td>0.06</td>
</tr>
<tr>
<td>Female 9</td>
<td>1</td>
<td>0</td>
<td>1.09</td>
<td>-0.09</td>
</tr>
<tr>
<td>Female 10</td>
<td>1</td>
<td>0</td>
<td>0.81</td>
<td>0.19</td>
</tr>
<tr>
<td>Male 1</td>
<td>0</td>
<td>1</td>
<td>0.40</td>
<td>0.60</td>
</tr>
<tr>
<td>Male 2</td>
<td>0</td>
<td>1</td>
<td>0.16</td>
<td>0.84</td>
</tr>
<tr>
<td>Male 3</td>
<td>0</td>
<td>1</td>
<td>-0.16</td>
<td>1.16</td>
</tr>
<tr>
<td>Male 4</td>
<td>0</td>
<td>1</td>
<td>0.05</td>
<td>0.95</td>
</tr>
<tr>
<td>Male 5</td>
<td>0</td>
<td>1</td>
<td>-0.01</td>
<td>1.01</td>
</tr>
<tr>
<td>Male 6</td>
<td>0</td>
<td>1</td>
<td>0.41</td>
<td>0.59</td>
</tr>
<tr>
<td>Male 7</td>
<td>0</td>
<td>1</td>
<td>0.04</td>
<td>0.96</td>
</tr>
<tr>
<td>Male 8</td>
<td>0</td>
<td>1</td>
<td>-0.02</td>
<td>1.02</td>
</tr>
<tr>
<td>Male 9</td>
<td>0</td>
<td>1</td>
<td>-0.03</td>
<td>1.03</td>
</tr>
<tr>
<td>Male 10</td>
<td>0</td>
<td>1</td>
<td>-0.05</td>
<td>1.05</td>
</tr>
</tbody>
</table>

PLS-DA of $X$ and $Y$ gave a one-component model. $R^2_X$, the variance in $X$ explained by the model, was 0.75. $R^2_Y$, the variance in $Y$ explained by the model, was 0.84 and $Q^2$ was 0.82 (The limit for a significant component was $Q^2 > 0.05$.). The $t[1]/t[2]$ score plot of the PLS-DA (Figure 1.8 (left)) is useful to overview the class discriminating ability of the model. To further interpret the model, the parameter variable influence on projection (VIP) can be inspected (Figure 1.8 (right)). VIP is a condensed summary of a PLS model, showing the influence of each $X$-variable on the model. In Figure 1.8 (right) the VIPs are sorted in descending order of importance and it can be concluded that the variable ’weight’ was the most important for the separation.

Class predictions can be performed by leave-one-out cross validation. The information in $Y$ is left out for one individual at a time while the model predicts the information in $Y$ based on the data in $X$. A cut-off of 0.5 means that a predicted $Y$ value below/above 0.5 is categorized as a correct prediction if the correct $Y$ value is zero/one. On the other hand, a predicted $Y$ value above/below 0.5 is categorized as an incorrect prediction if the correct $Y$ value
is below/above zero/one. Inspection of Table 1.3, containing class predictions based on the significant component, will probably clarify this. As shown in Table 1.3 we were very lucky; there were no misclassifications, and the accuracy of the binary classification test was 100%.

In study I-II PLS-DA of MRS data was applied to create predictive models that, using linear combinations of the metabolites, best separate groups within the data.

1.2.3 Independent components analysis

A commonly used set up in an fMRI experiment is to instruct the subject to perform experimental and control tasks in alternating sequence blocks. A reference function that contains information about how and when the tasks were carried out can then be constructed. Model based techniques such as SPM [23] use such a reference function to separate the signals of interest and the signals not of interest. The multivariate method independent component analysis (ICA) on the other hand allows extraction of signals of interest and not of interest without any prior information about the task. One example of a situation where that can be useful is resting state fMRI data. Calhoun et al. has proposed a model to apply ICA to group studies of fMRI data [24]. It is based on the assumption of statistical independence of the extracted component maps (‘spatial ICA’). The GIFT toolbox (http://icatb.sourceforge.net/groupica.htm) supports a group ICA approach, which first concatenates the individual data set from each subject, followed by the computation of the subject-specific ICA components and corresponding time courses. There are three main stages in group ICA of fMRI data: data compression (with PCA), application of the ICA algorithm and back reconstruction for each individual subject.
A number of different ICA algorithms are available. The majority of applications of ICA to fMRI use infomax [25], as the sources of interest are super-Gaussian in nature and the algorithm favors separation of super-Gaussian sources.

The idea behind ICA is that there are latent variables, sources, hidden in the data. It is assumed that they are statistically independent and that they are (linearly) mixed in the data. The aim of ICA is to recover the signals with minimum a priori information. The observed data, $x$, is modeled by a linear latent variable model

$$ x = As $$

where $A$, the mixing matrix, is constant (a parameter matrix) and $s$ contains the latent random variables called the independent components. The goal is to estimate $A$ and $s$, observing only $x$ under the assumptions that the sources, $s_i$, are mutually independent and non-Gaussian.

Each independent component provides a spatial distribution pattern in the imaged brain volume that share sufficiently similar temporal response pattern thus providing a natural measure of functional connectivity [26]. A component consists of a time course and a spatial map. Figure 1.9 presents an example of a component calculated with ICA of resting state fMRI data.

*Figure 1.9: Example of ICA component: spatial map and corresponding time course.*
1.3 Overview of investigated diseases and conditions

In this thesis MR experiments investigating the central nervous system (brain and spinal cord) were performed on both animals (rodents) and humans. Three different animal models were examined with proton MRS; a genetically modified mouse model for Alzheimer’s disease (AD), a rat model for spinal cord injury (SCI) and a genetically modified mouse model for premature aging. Also, human resting state fMRI data of young and old controls were acquired.

Translational researches are laboratory investigations targeting disease problems identified in the health care. In this type of research animal models for defined diseases in humans play an important role. The basic idea is that the results obtained in an investigated animal model can be generalized and extrapolated to humans. Genetically modified animals can serve as model systems of human diseases although an animal model rarely fully mirrors the human state in health or disease [27]. Instead an animal model might reflect one aspect of the investigated condition. One of the benefits of using animal models for scientific examinations rather than human subjects is the homogeneity of the animal group. In human experiments there might be variations within the group caused by e.g. age, diet, lifestyle, genes or diseases other than the investigated. In animal experiments the individuals are often identical twins, eating the same food and living in similar cages, which leads to a more homogenous group. From a statistical point of view this is preferable since the experimental variation of the investigated population is kept at a minimum.

1.3.1 Alzheimer’s disease

AD is a progressive neurodegenerative disorder that leads to dementia; a chronic, usually progressive disease first characterized by memory impairment and later by deterioration of intellectual capacity. AD is responsible for approximately two thirds of all dementia cases and is distinguished from frontotemporal dementia, Lewy body dementia and vascular dementia. AD is characterized by the destruction of nerve cells and neural connections in the cerebral cortex of the brain and by a significant loss of brain mass [28]. The primary criteria to diagnose AD relies on clinical observations, however a definite diagnosis is today only possible by a postmortem examination of brain tissues revealing amyloid plaques and neurofibrillary tangles. Amyloid plaques consist of deteriorating neuronal material surrounding deposits of a sticky protein, beta-amyloid [28]. Neurofibrillary tangles are twisted protein fibers located within nerve cells [28]. Early diagnosis of AD is difficult and the identification of novel early in vivo biomarkers is therefore of substantial interest. Neuroimaging has been used to complement clinical assessments in the early detection of AD.
Approximately 25% of all AD is familial (i.e. \( \geq 2 \) persons in a family have AD) of which approximately 5% is early-onset (age < 65 years) [29]. It is from these unusual early-onset forms of AD that genetic engineers have been able to transfer defect genes to laboratory animals and create transgenic animals that develop AD-like changes in the nervous system. Three forms of early-onset familial AD caused by mutations in one of three genes (APP, PS1, PS2) have been recognized. Different mouse models featuring a range of aspects of AD are available. Many models show different levels of beta-amyloid expression whereas other models also produce tangles. In this thesis a double transgenic mouse model, APP/PS1, was investigated.

1.3.2 Spinal cord injury

The spinal cord serves as a conduit for motor information, which travels down the spinal cord and for sensory information, which travels up the spinal cord. It also serves as a center for coordinating certain reflexes. If the spinal cord is injured the connection can be partly or completely cut off between the brain and the nerves in the spinal cord. Therefore, nerves from the brain can no longer give signals to the muscles and the information from nerves in the body can not reach the brain. SCI is classified as either complete or incomplete. A complete injury is indicated by a total lack of sensory and motor function below the level of injury. People with incomplete injuries retain some motor or sensory function below the injury. Every incomplete injury is unique and how much of the body that will be paralyzed depends on where the injury is, and how extended it is. If the spine is injured in the neck region, nerve paths to legs, trunk and arms will be affected (tetraplegia). If the injury is located in the thoracic spine or lower back, legs and parts of the trunk will be affected (paraplegia). In this thesis, data from rats with complete paraplegia were investigated.

1.3.3 Aging

Aging takes place at different levels in an organism, from DNA through cells, tissue, organs, organ systems, to the individual as a whole. A mitochondrial theory of aging [30] suggests that damage to mitochondrial DNA (mtDNA), mutations, slowly accumulates with time. One point mutation is a change within a gene in which one base pair in the DNA sequence is altered. The role of mtDNA point mutations in aging has been questioned, because the overall level of mtDNA mutations is usually lower than the threshold needed to cause respiratory chain dysfunction [31]. mtDNA polymerase is an enzyme that, among other things, repair mutations, i.e. it ‘proofreads’ the mtDNA. In this thesis we investigated mutator mice [32, 33] in which the proofreading ability of mtDNA polymerase is deficient. Therefore, mutations accumulate at a much higher rate than normal. There is a threefold to fivefold increase in
the levels of point mutations, as well as increased amounts of deleted mtDNA. The increase in mtDNA mutations is associated with reduced lifespan and premature onset of aging-related phenotypes.
2. Aim of thesis

MRS and fMRI can provide detailed information about complex metabolic activity in the central nervous system. The aim of this thesis was to explore how the information can be extracted and analyzed using model free and model driven MVDA.

Specific aims

*Study I*
To investigate how the neurochemical profile in mice, as measured by *in vivo* \(^1\)H MRS, changes during normal aging and in a model for AD. To test if, and at what time point, it was possible to differentiate wild type control mice from transgenic AD mice using MVDA of \(^1\)H MRS data.

*Study II*
A variety of tests of sensorimotor function are used to characterize outcome after experimental spinal cord injury. These tests typically do not provide information about chemical and metabolic processes in the injured central nervous system. In this study we wanted to investigate the potential of \(^1\)H MRS and MVDA for monitoring chemical changes in the central nervous system *in vivo* following spinal cord injury in rats.

*Study III*
To analyze the role of mitochondrial dysfunction and abnormal metabolism on central nervous system aging in wild type control mice and prematurely aging mtDNA mutator mice *in vivo* using \(^1\)H MRS and *ex vivo* using high-performance liquid chromatography, histology and biochemistry methodology.

*Study IV*
To investigate putative changes in functional connectivity networks associated with aging using ICA of human resting-state fMRI data. To evaluate the resting-state data with three different pre-processing procedures.
3. Materials and methods

The studies were performed with superconducting, horizontal bore magnets of magnetic field strength ranging from 3.0-9.4 T (see Table 3.1). Studies I-III were based on rodent in vivo $^1$H MRS data, while human resting state fMRI data were investigated in study IV.

**Figure 3.1:** Overview of study subjects, type of MR data generated and data processing in the thesis.

3.1 Magnetic resonance spectroscopy

In study I-III, MRS and MRI data were generated in vivo in rodents. The studies were approved by the Stockholm Ethics Committee and all experiments were performed in accordance with guidelines from the Swedish Animal Welfare Agency. Food and water were provided ad libitum and animals were kept on a 12/12 h light/dark cycle. Anesthesia was induced with isoflurane and maintained during scans by spontaneous breathing of about 2 % isoflurane. Body temperature and respiratory rate were monitored continuously. Reference images for positioning of the VOI were acquired using a spin echo sequence with rapid acquisition with relaxation enhancement (RARE) [34] in axial, sagittal and coronal planes. Effective TE was 25.16 ms in study I and
Table 3.1: Overview of hardware, magnets and coils used in this thesis together with their location in Sweden.

<table>
<thead>
<tr>
<th>Study</th>
<th>Location</th>
<th>Magnet</th>
<th>$B$ [T]</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>AstraZeneca (Södertälje)</td>
<td>Bruker Biospec Avance 94/30</td>
<td>9.4</td>
</tr>
<tr>
<td>II and III</td>
<td>Experimental MR research centre, Karolinska Institutet (Solna)</td>
<td>Bruker Biospec Avance 47/40</td>
<td>4.7</td>
</tr>
<tr>
<td>IV</td>
<td>Karolinska Sjukhuset (Huddinge)</td>
<td>Siemens Trio</td>
<td>3.0</td>
</tr>
</tbody>
</table>

37.41 ms in study II and III. Voxel shape and localization was achieved by PRESS [1] using Hermite radio frequency pulses with matched bandwidths. To achieve sufficient accuracy for quantification, a repetition time of sufficient length (3500 ms) was chosen, allowing complete relaxation of most metabolites in the spectrum between consecutive scans. Spectra were acquired with 256-1024 averages and water suppression (VAPOR [5]). Outer volume suppression was used to avoid spectral contamination.

Metabolites in the resulting spectra were quantified by the software package LCModel [3, 35]. Corresponding LCModel simulated basis set (provided by Stephen Provencher) matched $B_0$ field strength, localization sequence and TE. Phasing, referencing and quantitation is done automatically with LCModel. Metabolite concentrations were given relative to tCr. In study I and II, data were also analyzed with MVDA using [16]. Before MVDA, data were zero meaned and scaled to unit variance.

3.1.1 Study I

MR data were generated from brains of double transgenic APP/PS1 mice and wild type mice. Animals were investigated with in vivo $^1$H MRS, to examine the neurochemical profile of a VOI (8 mm$^3$) positioned in dorsal hippocampus (see Figure 1.1 on page 3). Animals were also examined with 3D MRI to enable volume measurements of anatomical regions. To study features of early AD, transgenic and wild type mice were investigated at the time points 2.5, 6.5 and 9 months of age, with the first time point chosen before first plaques appear. Histology of individual animals was performed (three transgenic animals sacrificed at each time point) to verify the absence or presence
of amyloid deposits. MR data were acquired using a 9.4 T Bruker magnet. A 72 mm volume coil was used for excitation and a quadrature mouse brain surface coil was used for signal detection. Localization was achieved by short (20 ms) TE PRESS. Fastmap was used for localized shimming. Metabolite concentrations were quantified relative to tCr and spectral data associated with CRLB > 50% were excluded from further analysis. MRS data were also analyzed with PLS-DA to investigate if transgenic mice could be distinguished from wild type mice. Leave-one-out cross-validation was applied, and sensitivity and specificity were calculated. VIP parameters of the metabolites were determined to examine their relative importance for the separation of groups. Manual volume segmentation were performed in the 3D MR images in which hippocampus and brain were outlined and lateral ventricles measured.

3.1.2 Study II

Rats with SCI were investigated using in vivo $^1$H MRS at several time points starting with naive control rats. Four different VOIs were investigated positioned in the cerebral cortex (2 VOIs, 72 mm$^3$ and 18 mm$^3$ respectively), thalamus/striatum (76 mm$^3$) and lumbar spinal cord (16 mm$^3$) (see Figure 4.2 on page 29). A bilateral cortex VOI was used to monitor short-term and long-term metabolic changes. A unilateral cortex VOI was centered over the sensorimotor area of the hind limbs, to investigate short-term changes. Another VOI, including thalamus/striatum, located in the deep centre of the brain, was used to examine long-term changes. Spectra were also acquired from a VOI in the lumbar spinal cord, beneath the injury, to monitor both short-term and long-term changes. MR data were acquired using a 4.7 T Bruker magnet. A linear bird cage resonator (Bruker, Ettlingen, Germany) with an inner diameter of 35 mm was used for excitation and reception to acquire bilateral cortex and thalamus/striatum spectra. A 72-mm bird cage resonator for transmission and
a ‘rat brain’ quadrature receiver coil (Bruker, Ettlingen, Germany) secured to the animal holder above the head was used to acquire the unilateral cortex VOI spectra. For acquisition of spinal cord spectra, a surface coil (T9510; Bruker, Ettlingen, Germany) with an inner diameter of 20 mm was used for transmission and detection. Localization was achieved by short TE PRESS. B₀ was optimized at the VOI using linear shims. Metabolite concentrations from all VOIs were quantified relative to tCr and spectral data associated with CRLBs > 20% were excluded from further analysis. Data for each metabolite were tested for normal distribution and for homogeneity of variances. The data were then analyzed VOI-wise by multivariate analysis of variance (MANOVA) followed by separate univariate analysis of variance (ANOVA) of the metabolites. Post hoc tests were applied to determine which groups (time points) differed. Data were also investigated with MVDA using PCA and PLS-DA. Leave-one-out cross-validation were applied, to estimate the overall predictive power of the models. VIP parameters were calculated.

3.1.3 Study III

Normally aging mice and prematurely aging mtDNA mutator mice [32, 33] were investigated in this study. Animals were examined with in vivo ¹H MRS in two VOIs located in cerebral cortex (10 mm³) and in striatum (20 mm³). MR data were acquired with the same magnet as in study II. A linear bird cage resonator (Bruker, Ettlingen, Germany) with an inner diameter of 25 mm was used for excitation and reception. Localization was achieved by short (16 ms) TE PRESS. B₀ was optimized at the VOI using linear shims. Metabolite concentrations were quantified relative to tCr and spectral data associated with CRLBs > 50% were excluded from further analysis. The data were then analyzed VOI-wise by ANOVA. mtDNA mutator mice and normally aging mice were further investigated ex vivo using high-performance liquid chromatography (HPLC), histology and biochemistry methodology.

3.2 Functional magnetic resonance imaging

3.2.1 Study IV

Human resting state fMRI data of two groups of healthy controls, young and elderly, were evaluated. The resting-state fMRI data were acquired using a Siemens whole-body 3T clinical MRI system (Magnetom Trio, Erlangen, Germany). Three different preprocessing pipelines were conducted (see Figure 3.3) in batch mode with shell scripts. ICA of the data were performed in batch mode using the program Group ICA of fMRI Toolbox (GIFT) version 1.3h. The approach is to first concatenate the individual data set from each subject, followed by the computation of
**Preprocessing of resting-state fMRI data**

- Exclusion of ten first timeframes
- Head motion correction
- Removal of scull
- Creation of whole-brain mask
- Spatial normalization to MNI template

**Set 1**

- Low-pass filtering and linear de-trending

**Set 2**

- Further de-trending by removal of polynomial up to the cubic order

**Set 3**

- Regression analysis with the following regressors:
  - Motion corrected parameters
  - Global signal based on whole-brain mask
  - CSF signal based on CSF mask
  - Average white matter signal based on white matter mask

---

**Figure 3.3:** Three different preprocessing pipelines in study IV.

The subject-specific independent components and corresponding time courses. Data reduction, using PCA, was performed both for individual subject data and group data followed by ICA (infomax) on the reduced data-set.

Aging related changes were statistically assessed by voxel-wise student t-test between subjects of the two age groups. A threshold at $t > 3.5$ and a minimum spatially connected cluster size $> 60$ voxels were employed. The inter-network coherence was evaluated by computing the cross-correlation between the time courses for each functional connectivity networks. The cross-correlation was evaluated for each dataset that underwent different pre-processing steps and for each individual subject of the different age groups. Histograms for the group t-maps were also computed for each independent component.
4. Results

4.1 Study I

In Figure 1.3 on page 4 and example spectrum from this study is shown. Shim- ming resulted in unsuppressed water signal line widths with full width at half maximum (FWHM) of 15-20 Hz. Fifteen metabolites were considered as de- tected and were further analyzed.

Group comparisons of individual metabolites revealed significant differ- ences between transgenic and wild type mice for Ins and Gua at 2.5 months age, and for Glu, NAA and macromolecules at 1.2 ppm at both 6.5 months and 9 months of age. PLS-DA of the quantified metabolites resulted in one or two significant ($Q^2 > 0.05$) components for the investigated time points. t[1]/t[2] scatter plots are presented in Figure 4.1, showing an increased separation between the transgenic and wild type mice as they grew older. Leave-one-out cross-validation with cut off value 0.5 classified individuals as transgenic or wild types with an accuracy of 80%, 88% and 100% at 2.5, 6.5 and 9 months of age respectively.

Volume measurements, based on 3D MRI, revealed that APP/PS1 mice had 5-6% smaller brains and 6-15% smaller hippocampus than wild type mice. Moreover, the areas measured over lateral ventricles were 20-33% larger in transgenic mice as compared with wild type mice. Histology showed that amy- loid plaques were present in mice at the age of 6.5 and 9 months, but not in 2.5 months old animals. No plaques were found in wild type animals (N=3) at 9 months of age.
Figure 4.1: Scatter plots of PLS-DA models of transgenic and wild type animals at the age of 2.5 (a), 6.5 (b) and 9 (c) months.
4.2 Study II

Figure 4.2 shows the VOI locations together with examples of corresponding \textit{in vivo} $^1$H MR spectra from the four investigated regions in the central nervous system. Shimming resulted in unsuppressed water signal line widths with FWHM of 8-16 Hz. Figure 4.3 presents an overview of the quantified metabolites from the four investigated VOIs. Concentration ratios of 4-8 metabolites were quantified in the VOIs in cortex, thalamus/striatum and the spinal cord of rats with SCI and control rats. Mean CRLBs corresponding to the quantified metabolites are also shown. Ins, tCho, tNAA and Glx were discernible in spectra from all the examined regions. In addition, Glu was detected in all brain spectra, and Gln in the bilateral VOIs, cortex and thalamus/striatum. Moreover, GABA was detected in thalamus/striatum and taurine in bilateral cortex.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig4_2}
\caption{Multislice RARE images of the rat brain with the volumes of interest (VOIs) bilaterally in cerebral cortex, thalamus/striatum, unilaterally in cerebral cortex and spinal cord and representative \textit{in vivo} $^1$H MR spectra measured from the corresponding brain regions.}
\end{figure}

The statistical distributions of data for each metabolite in all studied regions and at all time points were assessed to be normal. MANOVA revealed significant changes over time for the investigated metabolites for the bilateral cortex.
VOI and for the spinal cord VOI, but not for the regions investigated in unilateral cortex or in thalamus/striatum. Stars in Figure 4.3 indicate metabolites for which significant changes over time were found using ANOVA. Significant changes were found in the bilateral cortex VOI for Glu, tCho and Glx. In the spinal cord, significant changes were found for Ins, tCho, tNAA and Glx. No significant changes were found for any of the detected metabolites in the VOIs in unilateral cortex or in thalamus/striatum.

Figure 4.3: Mean concentration ratios of brain metabolites and corresponding CRLBs quantified by LCModel in bilateral cortex, unilateral cortex, thalamus/striatum and the spinal cord. Stars indicate metabolites for which significant differences were found over time.
4.3 Study III

Figure 4.4 presents spectra acquired in cerebral cortex and in striatum of mtDNA mutator mice and control mice. Shimming resulted in unsuppressed water signal line widths with FWHM of 8-13 Hz.

![Figure 4.4: Baseline corrected $^1$H MR spectra obtained from the VOIs in cerebral cortex and in striatum of a 33-week-old mutator mouse and an age-matched control.](image)

Lactate levels quantified by LCModel in the in vivo $^1$H MR spectra were higher in the mtDNA mutator mice compared with controls (Figure 4.5). In cerebral cortex, lactate levels were increased twofold in 6- to 9-week-old and threefold in 35- to 38-week-old mtDNA mice compared with wild type littermates. In striatum, lactate levels were also twofold higher in 6- to 9-week-old mtDNA mutator mice and these levels were maintained as mtDNA mutator mice aged. During normal aging lactate levels were found to be increased later in life. Mean CRLBs of lactate in mutator mice (10.7 (± 4.0 SD) in cortex and 7.8 (± 2.4 SD) in striatum) were lower than in control animals (26.5 (± 13.7 SD) in cortex and 12.2 (± 3.8 SD) in striatum).

Brain lactate levels measured in mtDNA mice and controls using HPLC confirmed the MRS findings. HPLC showed a threefold increase in brain lactate in 25-week-old mtDNA mutator mice, while the levels in wild type mice were within the physiological range as reported by others.

Further analysis ex vivo using histology and biochemistry led to the conclusion that high lactate levels in both brain and peripheral tissues (liver and heart) are the result of a metabolic shift to a glycolytic or anaerobic condition, where large amounts of lactate are being produced from pyruvate in an environment with increasingly dysfunctional mitochondria.
Figure 4.5: Two-way between subjects ANOVA on age-matched groups was performed on brain lactate concentrations (mean values ± SEM) in cerebral cortex (left) and striatum (right) from female mtDNA mutator (n = 19, red), compared to littermate wild-type (n = 18, black) mice. All mutator mice showed a significant increase in lactate levels and ANOVA demonstrated a significant effect for the mtDNA mutator allele both in cerebral cortex (F(1, 25) = 84.1, P < 0.0001) and striatum (F(1, 24) = 134, P < 0.0001), compared to controls. Post hoc analysis are denoted with *** P < 0.001.

4.4 Study IV

Based on t-score threshold and visual inspection of the t-maps and the corresponding time courses, twelve components were identified as relevant functional connectivity networks, eighteen components as artifacts related to CSF, vascular, susceptibility or motion-related artifacts and six components as mixed, representing brain functional networks contaminated by cerebrospinal fluid, motions and large veins. Figure 4.6 presents the twelve components common for the entire subject group (preprocessing Set 2) showing significant (p < 0.001) synchronized low-frequency fluctuations in the resting-state BOLD fMRI signal intensities. The observed networks are almost identical to those reported previously by Damoiseaux et al. [9]. In addition to the previously identified ten functional connectivity networks [9], we observed two more consistent functional networks in the studied subjects.

The results for the resting-state fMRI data that underwent different levels of preprocessing are quite consistent except for the increased anti-correlation with the removal of the global signal. Though the overall ICA results for the entire subject pool are largely comparable across the different stages of preprocessing, the results from band-pass filtering (Set 1) and higher order baseline correction (Set 2) were more similar to each other than those after the removal of the global, white matter, and CSF signals (Set 3). Figure 4.7 shows the t-map histogram comparisons between data that underwent different levels of pre-processing for the independent component corresponding to the default mode network. As shown, the removal of global, white matter, and CSF signals introduced more voxels with high negative t-scores (anti-correlation) in the network in addition to the overall statistical degradation for the estimation of the component, as demonstrated by increased number voxels with low t-scores and reduced number of voxels with high positive t-scores.
Figure 4.6: 12 relevant function connectivity networks identified by group ICA of the resting-state fMRI data acquired in the entire subject group after the second-stage pre-processing including only band-pass filtering and the third order baseline corrections.

The group comparison between elderly and young subjects revealed decreases in network coherence and connectivity with increasing age (Figure 4.8). The young subject group exhibited higher inter-network coherence than the elderly subject group.
Figure 4.7: Histograms for the t-maps of the independent component corresponding to the default mode. The red (Set 1), blue (Set 2) and black (Set 3) curves represent the results for the data sets that underwent different levels of pre-processing, respectively. The inserts are the amplifications for the tails of the distribution.

Figure 4.8: A summary of the brain regions that showed significant differences ($p < 0.01$ and cluster size $\geq 60$ voxels) in functional connectivity as identified by the ICA of the resting-state fMRI data between the healthy elderly subjects and young control groups.
5. Discussion

Study I

Transgenic APP/PS1 mice and control mice were investigated using $^1$H MRS. Metabolites were quantified relative to tCr with LCModel and further analyzed with PLS-DA.

NAA is found primarily in mature neurons and neuronal processes, such as axons. We found smaller amounts of NAA in the transgenic mice compared with wild types when the animals were 6.5 and 9 months old. At these time points the mice also showed plaque loads. Decreased levels of NAA in transgenic mice compared with wild types have previously been described by von Kienlin et al. [36] in the APP/PS2 mouse model at 24 months of age, by Marjanska et al. [37] in the APP/PS1 mouse model at 16 months of age and by Chen et al. [38] in the APP/PS1 mouse model at 8 months of age. von Kienlin et al. [36] investigated if the metabolite/tCr ratios could be correlated with plaque load (examined in histological data). They found a significant negative correlation of NAA with plaque load. In AD patients, NAA has been reported to decrease [39].

Glu is abundant in the central nervous system and acts as an excitatory neurotransmitter. We quantified smaller amounts of Glu in the transgenic mice compared with wild types at the age of 6.5 and 9 months of age. A decrease in the levels of Glu in transgenic mouse models for AD compared with wild type mice has been described earlier [36, 37].

Ins is a sugar alcohol with a structure similar to that of glucose, and is mainly considered as a glial marker. Ins has been found to increase in AD in humans. Reported results in the literature for Ins in mouse models for AD are diverse. We found less Ins in transgenic mice only at the first investigated time point. Marjanska et al. [37] investigated Ins levels in APP/PS1 and wild type mice 66-904 days of age. No increase in Ins levels were observed in animals between 66 and 400 days old. Marjanska et al. however found increased levels of Ins in APP/PS1 mice older than 400 days. Dedeoglu et al. [40] found no significant difference in m-Ins levels when investigating 10 and 12 months old transgenic APP and wild type mice. von Kienlin et al. [36] found no significant differences between transgenic mice and wild types, an no correlation between m-Ins and plaque load. Chen et al. [38] studied the APP/PS1 mouse model and found elevated levels of Ins/tCr in transgenic mice compared with wild types at 3, 5 and 8 months of age.
PLS-DA resulted in models that classified individual animals with increasing accuracy, 80% - 100% at 2.5 - 9 months of age. Lower levels of Ins and Gua in transgenic mice coincide with the first separation of transgenic and wild types. At this time point, volume defects in transgenic brains were seen. Later in life (6.5 and 9 months), when amyloid plaques were present, the separation between the groups became even stronger, and the involved metabolites were Glu, NAA and macromolecules at 1.2 ppm.

In conclusion, group differences in brain metabolites acquired in vivo with $^1$H MRS were found for APP/PS1 mice and wild type mice. First differences in metabolite content were seen at 2.5 months, when volume defects in transgenic mice were present, but no amyloid plaques. PLS-DA of MRS data showed that transgenic mice could be distinguished from wild type mice with 80% accuracy before plaques were formed.

**Study II**

In this study we compared acute and chronic consequences for the brain and lumbar spinal cord of low thoracic SCI in rats. Acutely in cortex decreased Glx levels were found, which slowly returned to normal levels. Conversely, this metabolite sum increased in the lumbar spinal cord, in which Glx levels remained increased 14 days post-injury, but were normalized 4 months after injury. In cortex, we were able to determine that the Glx decrease was caused mainly by a decrease in Glu. Additional changes of $^1$H-MRS-identifiable metabolites included alterations of tCho in the brain and spinal cord, and of Ins and tNAA in the spinal cord. The most conspicuous was a marked increase of Ins in the spinal cord below injury, seen first after 3 days and remaining at 4 months, the longest postoperative interval studied. Given the roles of Ins in intracellular signaling pathways, these observations point to marked and long-lasting alterations of cell signaling in spinal cord segments that no longer have bilateral axonal connections with the rest of the spinal cord and the brain. We found that thalamic and striatal tissue appeared less affected by the SCI, because no significant changes were identified in a VOI including these brain areas.

Qian et al. [41] used $^1$H-MRS to study rats with traumatically injured spinal cord and observed changes in NAA, tCr and tCho levels in segments caudal to the injury. In a study of $^1$H-MRS in SCI metabolic changes in thalamus/striatum of rats with SCI [42], MRS revealed an increase in the levels of NAA, Cr, Ins and Glu. In sham-operated animals increase in NAA and Cr levels were found.

The Glu/Gln/GABA cycle plays an important role in controlling levels of the major excitatory neurotransmitter Glu and the major inhibitory neurotransmitter GABA in the central nervous system. Gln produced in astrocytes is taken up by neurons, and converted to Glu or GABA. Glu released into the synaptic space is recycled by astrocytes. Excitotoxicity is thought to play a
role in the secondary degenerative events that follow SCI [43, 44, 45, 46], and increased levels of Glu have been found after SCI at levels neurotoxic even for neurons in the uninjured spinal cord [47]. To the extent that the marked alterations of Glx in the lumbar spinal cord may reflect alterations of Glu, as seems to be the case in cortex, the Glx increase in the spinal cord could help explain a state of hyperexcitability/increased reflexes often noted after SCI. In cortex, the loss of Glu presumably reflects decreased afferent activity, due to loss of input from hind limbs and other areas below the level of injury.

Ins has a key role in intracellular signalling and has also been regarded as an osmolite and glial marker [29, 48], and increases have been interpreted as increase of glial content or glial proliferation. After SCI, substantial astroglia occurs, and starting within 3 days after injury we found a significant increase of Ins in the spinal cord below injury, which was maintained throughout the studied postoperative time. This increase fits with the known long-term increase of glial fibrillar acidic protein-immunoreactivity in most parts of the spinal cord after injury. In group comparisons of spinal cord spectra, Ins was clearly the most important metabolite at 3 days and at all later time points. In the brain, Ins levels were fairly stable, and did not change significantly over time. This is in line with the lack of any marked astroglial responses in the brain to SCI.

GPC and PCh are the main components of the prominent choline peak, and constitute metabolites of phospholipid components of cellular membranes. Compared with the changes of Ins, we found only moderate changes of tCho. We noted a 16 % decrease of tCho in the spinal cord already 1 day after injury, a time when there was no change of Ins. This suggests that cell membrane changes occur faster in the spinal cord below injury than hitherto appreciated, and that astroglial changes occur as a consequence of the massive axonal damage.

NAA is seen as a prominent peak in the MRS, making NAA one of the most reliable markers for brain MRS studies. Under normal conditions, NAA is synthesized in and exported from the mitochondria, predominantly in neurons, and hence considered a neuronal marker for many brain diseases [49, 50, 51]. NAA increases are seen during development [52]. In a study of patients with incomplete SCI, NAA elevations were detected in the cerebral cortex [53]. Chromatography-mass spectrometry has also been used to study NAA concentrations up to 1 week after SCI in rats [54]. Caudal to injury, NAA levels were virtually indistinguishable from those in control animals. We found an increase of NAA in the lumbar enlargement of the spinal cord following SCI. The increase in the spinal cord was temporary, and might reflect a compensatory up-regulation, as NAA is considered to be an important osmolytic regulator for neurons, and/or a regulator of local sprouting. In humans with incomplete SCI [53], 50 % higher NAA levels in cerebral cortex were reported 0.5-2 years after injury.
Limitations of the technique as performed in the rat include the rather inhomogeneous VOI content in the small rodent central nervous system. Also, we had to scan five segments distal to the spinal cord lesion, in order for the surface coil to be close enough to cord tissue.

In study II, the initial idea was to simultaneously monitor metabolic changes in brain and spinal cord and put all the data into one model. However, it was not successful to build robust PLS-DA models for classifications in which data from different VOIs were combined. The PCA t[1]/t[2] score plot based on data from all investigated VOIs revealed groupings among the observations. The data separated into four clusters that can be interpreted as differences between the four selected central nervous system areas, but also between the three brain areas on the one hand and spinal cord data on the other, between bilateral and unilateral VOIs, and/or between the different coils that were used in this study. These distinctions overshadowed the differences between classes of animals at different time points. Models for classification were therefore built VOI-wise.

In conclusion, changes of metabolites in brain and spinal cord after SCI were found using in vivo ¹H MRS. Both long-term and short-term changes were investigated. Metabolite alterations were found distant to the site of primary damage, in bilateral cortex. In two other investigated brain areas, unilateral cortex and thalamus/ striatum, no changes in metabolite concentrations were found. Four metabolites, Ins, tCho, tNAA and Glx, were detected in the spinal cord caudal to injury. Changes over time were found for all detected metabolite ratios in the spinal cord.

Study III
Cerebral lactate metabolism and its compartmentalization in astrocytes, neurons, and elsewhere is not fully understood [55]. Lactate is continuously produced in brain, heart, skeletal muscle, and other tissues, even during completely aerobic conditions [56]. It has been suggested that lactate constitutes an alternative source of energy that the brain uses under strenuous situations [57]. It has been shown that, under conditions of increased lactate production (i.e. exercise), the use of blood lactate as an energy source in the brain increases at the expense of blood glucose [58]. Lactate is a substrate for the mitochondrial TCA cycle, and its oxidation can produce a significant amount of ATP [59].

When healthy aging in humans was recently accessed with combined ¹³C−/¹H−MRS, an association was found between reduced neuronal mitochondrial metabolism and altered glial mitochondrial metabolism in aged (76 ± 8 years) participants [60]. Another study found that lactate levels measured by ¹H-NMR in 88- to 96-weeks-old rats were significantly increased [61].
Comparing normally aging and mtDNA mutator mice allows us to conclude that the increased LDH-A/LDH-B gene expression ratio is causative of high brain lactate levels and that these lactate levels could predict aging. We have strong evidence that lactate levels are elevated in advance of other indices of aging in the prematurely aging mtDNA mutator mouse. We found that mitochondrial dysfunction in brain leads to a metabolic shift from aerobic respiration to glycolytic metabolism, resulting in expression changes of the lactate dehydrogenase genes (LDH-A, LDH-B). This shift results in increased brain lactate levels, detectable using $^1$H MRS, prior to the appearance of overt aging phenotypes.

**Study IV**

**Consistent functional connectivity networks**

The concept of resting-state functional connectivity suggests that the brain is spontaneously active in the absence of a goal-driven task, showing rich intrinsic dynamics, which can be modulated by external stimuli. As found in the present study, multiple previous resting-state fMRI studies [8, 9, 10, 62] have reported apparent inter-subject similarity in the identified network patterns. Damoiseaux et al. [9] quantitatively evaluated the inter-subject consistency of these resting-state network patterns. Reproducible network patterns consistent across subjects and sessions were found, and also the voxel-wise cross-subject variation for these networks. More recent studies [8, 62] based on much larger subjects pool further confirmed these findings.

**Preprocessing methods on the accuracy and reliability of group ICA results**

ICA has emerged as a robust technique to process resting-state and task-modulated fMRI data and to identify brain functional networks without detailed hypothesis about brain activations. Despite its widely application, there is little consensus on how data should be pre-processed prior to ICA. Here, we investigated the effects of three frequently used methods: 1) frequency filtering, 2) baseline correction (involves voxel-wise division of the time series mean and higher order de-trending) and 3) global signal removal. Band-pass filtering reduce contaminations from physiological artifacts associated with respiration and cardiac cycles [63]. Both baseline corrections and global signal removal aim to make distinction between global effects and the global signal. Global effects generally confound local signals in BOLD fMRI studies. They may reflect diffuse physiological processes or variations in scanner sensitivity and are difficult to measure directly. Particularly, in resting-state fMRI studies, the status of the resting-state is not so well defined. It is therefore necessary to understand the consequence of each pre-processing procedure. Improved auxiliary monitoring of the physiological activities, attention, alertness, and other mental activities using simultaneous EEG recordings during...
resting-state fMRI may provide bases for some specific data pre-processing procedures.

**Deactivation, anti-correlation and default mode network**

The concept of deactivation can be divided into task-independent and task-specific. Certain brain regions show a decreased BOLD activity during a variety of attention-demanding tasks in comparison to the resting-state baseline [12]. The baseline activity of this brain network has been defined as an organized default mode of brain function [12] that is suspended during task- or goal-directed brain activity. It was suggested that with the increasing workload in cognitive tasks, resources are redirected from the default mode network to task-specific cortical regions, resulting in decreased activity of the default mode network. Activity within the default mode network was hypothesized to be mostly inwardly directed high-order cognitive processes, e.g., goal-oriented planning, encoding, and memory functions [10, 64]. Task specific deactivations in normal controls have been a subject of extensive studies [65, 66, 67, 68, 69, 70, 71, 72]. Kawashima et al. [73], using a selective attention task, first demonstrated the existence of deactivation in brain irrelevant to the stimuli. It was suggested that the task-specific deactivation has the function to facilitate the task-specific activations through the suppression of task irrelevant cortical regions to enable the subject to focus the attention on the relevant task. However, another PET study by Friston et al. [74] demonstrated that the global flow was related with the experimental conditions, such as the magnitude of the adjusted local effect. Similarly, Aguirre et al. [75] reported that there was a significant correlation between observed global fMRI signals and an experimental paradigm. These results seem to suggest that global neuroimaging signals can be correlated with the experimental manipulations and are thus not necessarily simple nuisance variables to be excluded. The implication is that excluding the global signal variation in PET and fMRI analyses may not be simply increasing the statistical power, but meaningfully changing the results and hence interpretation of these studies. For example, the removal of global signal by linear regression in resting-state fMRI can mandate the introduction of anti-correlation into the identified functional connectivity networks, as demonstrated mathematically [76] and experimentally.

**Age effects on resting-state functional connectivity**

Both previous [8, 62] and the current study showed that normal brain aging can lead to extensive changes in functional connectivity and coherence. These changes do not seem to be very specific and are spatially widely distributed in a large number of functional networks. A relevant question for using the resting-state fMRI approach to assess aging is how these the resting-state functional networks in general, and the default mode network in particular, are affected by ongoing neuronal degeneration in elder subjects. Recently, abnormal resting-state functional connectivity patterns have been reported in differ-
ent brain regions in individuals at the risk for AD [77, 78, 79, 80, 81, 82, 83]. The involved brain regions include anterior prefrontal cortex, middle temporal lobe, posterior cingulate cortex, precuneus and parietal lobe. The most consistent observation is the decreased connectivity in posterior cingulate cortex, precuneus and prefrontal cortices, which are the important parts of the default mode network.
6. Acknowledgements


Tie-Qiang Li, tack för att jag haft förmånen att få jobba med dig. Du har imponerande kunskaper om MR och jag har försökt tillgodogöra mig så mycket jag kan av dem.


Rouslan Sitnikov, tack för den tid och det engagemang, som du satsat i den här avhandlingen och för att du både förklarade och problematiserade många aspekter av MR. Du introducerade mig till ett nytt sätt att se på MR och forskning.

Stefan Brené, tack för att du alltid har en positiv attityd, även när det ser mörkt ut.

Christian Spenger, tack för att du beredde väg för mig att bli doktorand vid Karolinska Institutet efter avslutat examensarbete på AstraZeneca.


Marianne Lundmark, varmt tack för goda råd och coachning, som gjort det möjligt för mig att slutföra arbetet avhandlingen.


Finansiell support gavs via KID-medel från Karolinska Institutet, Stiftelsen för gamla tjänarinnor samt Gun och Bertil Stohnes stiftelse. Denna forskning har använt sig av det medicinska bildbehandlingslaboratoriet SMILE vid Karolinska Universitetssjukhuset, Stockholm.
Bibliography


[16] SIMCA; Umetrics AB, Umeå, Sweden.


